CLAIMS

1. A method for identifying a subset of genes, comprising:
   identifying a first reference set of expressed genes, said first reference set consisting of genes differentially expressed between a first sample and a second sample; wherein said first and second samples differ with respect to a phenotype; identifying a second reference set of expressed genes, said second reference set consisting of genes that are differentially expressed between a third samples and a fourth sample; wherein said third and fourth differ with respect to said phenotype; identifying a concordance set of expressed genes, said concordance set consisting of genes common to said first and second reference sets wherein the direction of said differential expression is the same in said first and second reference sets; and identifying a subset of genes within said concordance set, wherein said subset is selected so that a first correlation coefficient, correlating for said genes within said subset a first expression differential between said first and second samples to a second expression differential between third and fourth samples, exceeds a predetermined value.
2. The method of claim 1, wherein said first correlation coefficient is selected from the group consisting of a correlation coefficient $\rho_{x,y}$, a Pearson product moment correlation, and a square of a Pearson product moment correlation coefficient.
3. The method of claim 1, wherein said differentials are logarithmically transformed prior to calculating said first correlation coefficient.
4. The method of claim 3, wherein said first correlation coefficient has an absolute value ≥ 0.8.
5. The method of claim 4, wherein said first correlation coefficient has an absolute value ≥ 0.9.
6. The method of claim 5, wherein said first correlation coefficient has an absolute value ≥ 0.95.
7. The method of claim 6, wherein said first correlation coefficient has an absolute value ≥ 0.995.

8. The method of claim 1, wherein said gene expression data from either or both of said first reference set and said second reference set is independently selected from the group consisting of mRNA quantification data, cRNA quantification data, cDNA quantification data, and protein quantification data.

9. The method of claim 1, wherein at least one of said first sample and said second sample comprises a cell line.

10. The method of claim 9, wherein said cell line is selected from the group consisting of a tumor cell line, a pluripotent precursor cell line, an omnipotent stem cell line, and a differentiated cell line.

11. The method of claim 10, wherein said cell line is a tumor cell line.

12. The method of claim 10, wherein said cell line is a pluripotent precursor cell line.

13. The method of claim 10, wherein said cell line is an omnipotent stem cell line.

14. The method of claim 9, wherein said first sample comprises a cell recovered from an orthotopic implant.

15. The method of claim 14, wherein said second sample comprises a cell recovered from an ectopic implant.

16. The method of claim 9, wherein at least one of said third sample and said fourth sample comprises a cell recovered from a patient.

17. The method of claim 9, wherein at least one of said third sample and said fourth sample comprises a cell recovered from a healthy donor.

18. The method of claim 16, wherein said cell is a tumor cell.

19. The method of claim 18, wherein said tumor cell is recovered from an organ selected from the group consisting of a prostate, a breast, a colon, a lung and an ovary.

20. The method of claim 1, wherein said phenotype is selected from the group consisting of recurrence, non-recurrence, invasiveness, non-invasiveness, metastatic, localized, tumor grade, Gleason score, survival prognosis, lymph node

status, tumor stage, degree of differentiation, age, hormone receptor status, PSA level, histologic type, and disease free survival.

21. The method of claim 1, wherein any of the group consisting of said first sample, said second sample, said third sample, and said fourth sample comprises a plurality of independent samples, and at least one of said first and said second differential is an average over said plurality of independent samples.

22. A method of correlating gene expression with a sample phenotype, comprising:
identifying a subset of genes according to the method of claim 1;
and determining the sign of a second correlation coefficient, said second correlation coefficient correlating for said genes within said subset said first or said second expression differential to an expression differential obtained from an unclassified sample, whereby the sign of said second correlation coefficient establishes a positive or a negative correlation with said phenotype of claim 1.

23. The method of claim 22, further comprising determining the magnitude of said second correlation coefficient and using said magnitude to assess the reliability of said established correlation.

24. The method of claim 22, wherein said subset consists essentially of the genes identified in Table 5, Table 7, Table 8, Table 9, Table 10, Table 13, Table 14, Table 15, Table 16, Table 18, Table 19, Table 20, Table 21, Table 22, Table 24, Table 25, Table 26, Table 27, Table 28, Table 29, Table 30, Table 31, Table 32, Table 33, Table 34, Table 35, Table 36, Table 37, Table 38, Table 41, Table 43, Table 44, Table 45, Table 46, Table 49, Table 50, Table 51, Table 52, Table 53, Table 55, Table 56, Table 57, Table 58, Table 61, Table 62, Table 65, Table 66, Table 67, Table 68, Table 69, Table 73, or Table 75.

25. The method of claim 24, wherein said subset consists essentially of 90% of the genes identified in Table 5, Table 7, Table 8, Table 9, Table 10, Table 13, Table 14, Table 15, Table 16, Table 18, Table 19, Table 20, Table 21, Table 22, Table 24, Table 25, Table 26, Table 27, Table 28, Table 29, Table 30, Table 31, Table 32, Table 33, Table 34, Table 35, Table 36, Table 37, Table 38, Table 41, Table 43, Table 44, Table 45, Table 46, Table 49, Table 50, Table 51, Table 52, Table

53, Table 55, Table 56, Table 57, Table 58, Table 61, Table 62, Table 65, Table 66, Table 67, Table 68, Table 69, Table 73, or Table 75.

26. The method of claim 25, wherein said subset consists essentially of 80% of the genes identified in Table 5, Table 7, Table 8, Table 9, Table 10, Table 13, Table 14, Table 15, Table 16, Table 18, Table 19, Table 20, Table 21, Table 22, Table 24, Table 25, Table 26, Table 27, Table 28, Table 29, Table 30, Table 31, Table 32, Table 33, Table 34, Table 35, Table 36, Table 37, Table 38, Table 41, Table 43, Table 44, Table 45, Table 46, Table 49, Table 50, Table 51, Table 52, Table 53, Table 55, Table 56, Table 57, Table 58, Table 61, Table 62, Table 65, Table 66, Table 67, Table 68, Table 69, Table 73, or Table 75.

27. The method of claim 26, wherein the subset consists essentially of 70% of the genes identified in Table 5, Table 7, Table 8, Table 9, Table 10, Table 13, Table 14, Table 15, Table 16, Table 18, Table 19, Table 20, Table 21, Table 22, Table 24, Table 25, Table 26, Table 27, Table 28, Table 29, Table 30, Table 31, Table 32, Table 33, Table 34, Table 35, Table 36, Table 37, Table 38, Table 41, Table 43, Table 44, Table 45, Table 46, Table 49, Table 50, Table 51, Table 52, Table 53, Table 55, Table 56, Table 57, Table 58, Table 61, Table 62, Table 65, Table 66, Table 67, Table 68, Table 69, Table 73, or Table 75.

28. The method of claim 27, wherein the subset consists essentially of 60% of the genes identified in Table 5, Table 7, Table 8, Table 9, Table 10, Table 13, Table 14, Table 15, Table 16, Table 18, Table 19, Table 20, Table 21, Table 22, Table 24, Table 25, Table 26, Table 27, Table 28, Table 29, Table 30, Table 31, Table 32, Table 33, Table 34, Table 35, Table 36, Table 37, Table 38, Table 41, Table 43, Table 44, Table 45, Table 46, Table 49, Table 50, Table 51, Table 52, Table 53, Table 55, Table 56, Table 57, Table 58, Table 61, Table 62, Table 65, Table 66, Table 67, Table 68, Table 69, Table 73, or Table 75.

29. A kit comprising a set of reagents useful for determining the expression of a subset of genes, said subset consisting essentially of the genes identified in Table 5, Table 7, Table 8, Table 9, Table 10, Table 13, Table 14, Table 15, Table 16, Table 18, Table 19, Table 20, Table 21, Table 22, Table 24, Table 25, Table 26,

Table 27, Table 28, Table 29, Table 30, Table 31, Table 32, Table 33, Table 34, Table 35, Table 36, Table 37, Table 38, Table 41, Table 43, Table 44, Table 45, Table 46, Table 49, Table 50, Table 51, Table 52, Table 53, Table 55, Table 56, Table 57, Table 58, Table 61, Table 62, Table 65, Table 66, Table 67, Table 68, Table 69, Table 73, or Table 75, and instructions for use.

30. The kit of claim 29, wherein the subset consists essentially of 90% of the genes identified in Table 5, Table 7, Table 8, Table 9, Table 10, Table 13, Table 14, Table 15, Table 16, Table 18, Table 19, Table 20, Table 21, Table 22, Table 24, Table 25, Table 26, Table 27, Table 28, Table 29, Table 30, Table 31, Table 32, Table 33, Table 34, Table 35, Table 36, Table 37, Table 38, Table 41, Table 43, Table 44, Table 45, Table 46, Table 49, Table 50, Table 51, Table 52, Table 53, Table 55, Table 56, Table 57, Table 58, Table 61, Table 62, Table 65, Table 66, Table 67, Table 68, Table 69, Table 73, or Table 75.

31. The kit of claim 30, wherein the subset consists essentially of 80% of the genes identified in Table 5, Table 7, Table 8, Table 9, Table 10, Table 13, Table 14, Table 15, Table 16, Table 18, Table 19, Table 20, Table 21, Table 22, Table 24, Table 25, Table 26, Table 27, Table 28, Table 29, Table 30, Table 31, Table 32, Table 33, Table 34, Table 35, Table 36, Table 37, Table 38, Table 41, Table 43, Table 44, Table 45, Table 46, Table 49, Table 50, Table 51, Table 52, Table 53, Table 55, Table 56, Table 57, Table 58, Table 61, Table 62, Table 65, Table 66, Table 67, Table 68, Table 69, Table 73, or Table 75.

32. The kit of claim 31, wherein the subset consists essentially of 70% of the genes identified in Table 5, Table 7, Table 8, Table 9, Table 10, Table 13, Table 14, Table 15, Table 16, Table 18, Table 19, Table 20, Table 21, Table 22, Table 24, Table 25, Table 26, Table 27, Table 28, Table 29, Table 30, Table 31, Table 32, Table 33, Table 34, Table 35, Table 36, Table 37, Table 38, Table 41, Table 43, Table 44, Table 45, Table 46, Table 49, Table 50, Table 51, Table 52, Table 53, Table 55, Table 56, Table 57, Table 58, Table 61, Table 62, Table 65, Table 66, Table 67, Table 68, Table 69, Table 73, or Table 75.

33. The kit of claim 32, wherein the subset consists essentially of 60% of the genes identified in Table 5, Table 7, Table 8, Table 9, Table 10, Table 13, Table 14, Table 15, Table 16, Table 18, Table 19, Table 20, Table 21, Table 22, Table 24, Table 25, Table 26, Table 27, Table 28, Table 29, Table 30, Table 31, Table 32, Table 33, Table 34, Table 35, Table 36, Table 37, Table 38, Table 41, Table 43, Table 44, Table 45, Table 46, Table 49, Table 50, Table 51, Table 52, Table 53, Table 55, Table 56, Table 57, Table 58, Table 61, Table 62, Table 65, Table 66, Table 67, Table 68, Table 69, Table 73, or Table 75.

34. The kit of any one of claims 29 -- 33, wherein said reagents are affixed to a solid support.

35. The kit of any one of claims 29 -- 33, wherein said reagents comprise primers for a nucleic acid amplification reaction.